SUPPLEMENTARY METHODS

To request access to B-SNIP data used in this manuscript, visit https://nda.nih.gov/. There is a "Get Data" tab at the top of the main page, under which is found a "Request Data Access" link. Request psychosis and healthy subject data, and the below described biomarker data, from:

B-SNIP1 (https://nda.nih.gov/edit_collection.html?id=2274)

B-SNIP2 (https://nda.nih.gov/edit_collection.html?id=2165)

PARDIP (https://nda.nih.gov/edit_collection.html?id=2126)

Data collection strategies were the same for B-SNIP1 [1-5] and the replication samples [6-9]. Data analyses were also the same for all biomarkers. In some instances, this mean re-scoring B-SNIP1 data to match updated procedures. Any differences between B-SNIP1 and replication samples are identified in what follows by underlined text, with a brief explanation.

Recruitment

B-SNIP recruitment sites were in Athens GA (replication sample only), Baltimore MD (B-SNIP1 only), Boston MA, Chicago IL (University of Illinois-Chicago for B-SNIP1 and University of Chicago for the replication sample), Dallas TX, Detroit MI (B-SNIP1 only), and Hartford CT. Each site recruited cases with possible with schizophrenia, schizoaffective disorder, or bipolar disorder with psychosis, and healthy persons for comparison. Psychosis cases were clinically stable and in a non- acute symptom state. The broad geographical span of B-SNIP facilitated enrichment of the study group by local geographical characteristics. Sites used a combination of newspaper, community, and mental health treatment facility advertising, with the groups similarly recruited across all sites. These subjects were a research sample; nonetheless, the large study numbers and broad geographical

recruitment enhance the generalizability of data from the B-SNIP cohort. This strategy generated a more inclusive study group than is typical in studies focusing on specific disorders, with the aim of having a representative sample of the spectrum of psychosis. Psychosis cases were limited to schizophrenia, schizoaffective disorder, and bipolar disorder with psychosis because these are the diagnoses with the highest prevalence of psychosis and studying more diagnostic categories was deemed unfeasible as a first approach.

Participants were assessed phenomenologically as described in the main text. The extensive clinical information on each participant was reviewed in a best-estimate diagnostic meeting with at least two experienced research clinicians to establish the consensus diagnosis. Cross-site diagnostic conference calls were carried out monthly; they were chaired by two senior primary investigators and attended by the 2–4 trained clinical assessors at each site. At study start, there was a face-to-face training session for all raters, with a requirement for reliability above 0.85. Each month, diagnostic conferences were held with in-depth diagnostic discussions. Each year, rater training was repeated to reestablish reliability. See Tamminga et al. [10] for complete details.

Procedures and Analyses

Following verification of study eligibility, subjects that met study inclusion criteria were scheduled for laboratory biomarker testing. Testing took place over 2-3 days at the recruiting sites. Recording and testing conditions were equivalent and stimulus presentation and recording equipment were identical across sites. Experimenters across sites also were trained and monitored to ensure identical laboratory data collection procedures across sites. As a result, there were no site effects that influenced group comparisons on any laboratory biomarker measure for either B-SNIP1 or replication samples.

Laboratory tasks for Biotype determination

The BACS and EEG/ERP procedures described here were updated from our original Biotypes paper [11]. All data from B-SNIP1 and the replication samples were scored using equivalent procedures as described below and as detailed in replication sample publications [6-9, 12]. Age and sex-adjusted data were used for all biomarkers based on the procedure described in Dukart et al [13].

Brief Assessment of Cognition in Schizophrenia (BACS)

The BACS is a 30 min, reliable, valid, and widely used test of neuropsychological function [6, 14-16]. It has six subtests covering four cognitive domains (Verbal Memory, Processing Speed, Reasoning and Problem Solving, Working Memory), although a composite score averaged over these four domains that assesses overall neuropsychological functioning yields an effective measure of psychosis-related cognitive performance [1, 17]. Rather than the composite measure that was used in the original Biotypes paper [11], we used principal component analysis (PCA: Covariance Matrix; Promax Rotation; Kappa 4) to integrate over the six subtests and create the BACS bio-factor. This made bio-factor formulation consistent across all measures.

Pro- and anti-saccade tasks

Participants performed pro- and anti-saccade tasks under identical conditions across sites and projects [2, 7]. Pupil position was recorded using EyeLink II head-mounted infrared headsets (500 Hz sampling rate) and their corresponding control platform (SR Research Ltd., Mississauga, Canada). Stimuli were programmed using Presentation software (Neurobehavioral Systems, Inc., Berkeley, CA,) and presented on 22-inch CRT monitors in completely darkened rooms.

Trials began with a red crosshair at screen center for visual fixation, followed by a white peripheral cue (+/-10 and 15 deg) with variable ITIs (1500-2500 msec). For prosaccade trials, the

peripheral cue was a white round dot (1-deg visual angle). Three fixation conditions were administered that altered the timing of fixation extinction relative to the peripheral illumination (32 trials per condition). For the "gap" condition, the central target was turned off 200 msec before the illumination of the peripheral cue. For the "synchronous" condition, fixation was turned off simultaneously with illumination of the peripheral cue. In the "overlap" condition, fixation remained on for 200 msec following illumination of the peripheral cue. Participants were instructed to fixate on the central cross and then to move their eyes as quickly and accurately as possible to the peripheral cue once it appeared. For antisaccade trials, the peripheral cue was a white square (1-deg visual angle). For antisaccade "overlap" condition (80 total trials), fixation and peripheral cue overlapped by 200msec. The overlap condition was used because it is most sensitive as an endophenotype in psychosis [18]. Participants were instructed to fixate on the central crosshair and then when the peripheral cue appeared, to move their eyes quickly and accurately to the mirror image location of the cue (opposite direction, same distance from central fixation). For the replication sample, there was also an anti-saccade gap condition [7], but those trials were not used here to maintain consistency with B-SNIP1.

The task order was always prosaccade tasks first, followed by antisaccade-overlap. Order of prosaccade tasks was counterbalanced between participants. Within each task and condition, trials were arranged pseudo-randomly so that trials were evenly split between +/-10- and 15-deg displacements. A brief practice block was performed before antisaccade-overlap to ensure participants' understanding of the antisaccade task.

Eye position data over time was scored by trained research assistants blind to participant group membership using in-house programs [7] developed in Matlab (MathWorks Inc., Natick, MA, USA). Trials characterized by anticipatory movements, blinks during cue onset, lack of saccades, or eye drifts were excluded. For each saccade the following variables were scored for (i) direction (for

evaluation of correct or error response) and (ii) onset latency (time from illumination of the cue to start of saccade). We used PCA (Covariance Matrix; Promax Rotation; Kappa 4) to integrate over saccade variables that significantly differentiated groups, as described in the results section, to create the saccade bio-factor scores.

Stop Signal task (SST)

All trials began with the subject seated before a computer monitor displaying a white central fixation cross for a 750-1500 msec. A green circle (the Go cue) then appeared to the left or right for 650 msec. On 40% of trials, a Stop Signal (red stop sign) was presented at central fixation at delays varying between 50-282 msec after the Go stimulus [3, 6]. Participants were instructed to respond as quickly and accurately as possible by pressing a button with their left index finger for stimuli appearing on the left side of the screen or another button with their right index finger for stimuli appearing on the right side. To maintain prompt responding, when participants did not respond within 650 msec on Go trials the trial was terminated and a red 'X' and the word "faster" were presented (these trials were not included in analyses; a proportion of them were re-presented; [3, 6]). On Stop trials in which the participants pressed a button and thus failed to inhibit the Go response, a red 'X' appeared over the stop sign to provide performance feedback, and these trials were counted as errors (see [3, 6] for additional details). A baseline task consisting of 50 consecutive Go trials, evenly and randomly distributed to cues on the left and right side of the screen, was administered to assess baseline reaction time to Go cues. Strategic slowing (difference between response latencies on baseline Go trials and Go trials during Stop Signal performance) and proportion of Stop Signal errors were used in Biotype construction [11]. PCA (Covariance Matrix; Promax Rotation; Kappa 4) integrated over the two SST variables, as described in the results section, to create the SST bio-factor score.

EEG Recording

Recording EEG were continuously recorded from 64 Ag/AgCl sensors (impedance < 5KΩ; Quik-Cap, Compumedics Neuroscan, El Paso, TX), positioned according to the standard 10-10 EEG system plus mastoids and CB1/2 locations to provide sampling below the canthomeatal line, with nose reference and forehead ground. Recordings were amplified (×12,500) and digitized (1000 Hz) using Neuroscan Acquire and Synamps2 recording systems (Compumedics Neuroscan).

EEG Data Pre-Processing

Raw EEG data were inspected for bad sensors and artifacts. Bad sensors were interpolated (<5% of all sensors for any subject) using spherical spline interpolation (BESA 5.3; MEGIS Software, Grafelfing, Germany). Data were converted to an average reference and digitally band-pass filtered from 0.5–55 Hz (zero phase filter; rolloff: 6 and 48 dB/octave, respectively). Blink and cardiac artifacts identified using independent components analysis were removed (EEGLAB 9.0; [19]).

Auditory Paired Stimuli Task

Psychosis and healthy analyses generally followed procedures established in Hamm et al. [5] and Clementz et al. [11] and were applied to all subjects in this paper. Minor modifications from B-SNIP1 publications ensured standardized data quality control between projects [8]. Changes to the scoring procedures were primarily to frequency domain quantification and were made to simplify the analysis steps to improve the ease of replication and verification by other laboratories. The procedures described in Parker et al. [8] were used for this report.

Stimuli. Recording conditions were equivalent and stimulus presentation and recording equipment identical across sites. While seated in a sound and electrically shielded booth (ambient sound = 61–63 dB; luminance = 0.11–0.12 foot-candles), subjects passively listened to 120-150 (B-

SNIP1: 150, Replication Sample: 120) binaural broadband auditory click pairs (4 msec duration at 75 dB sound pressure level; 500 msec inter-click interval) occurring an average of every 9.5 sec (9–10 sec inter-pair interval) and delivered through headphones. Participants who were smokers refrained from smoking 1 hr prior to testing.

Spatiotemporal Data Reduction. Data were segmented into epochs from 100 msec before to 850 msec after click-pair onset. The 100 msec pre-S1 period served as baseline. Epochs containing activity ±75 μV were eliminated. Data from good trials were averaged across trial types within a subject to create 64-sensor event-related potentials (ERPs).

In order to maximize use of available spatial, temporal, and oscillatory information in the evoked auditory response, a frequency-wise PCA (Covariance matrix, Promax Rotation, Kappa 3 with Kaiser normalization) of evoked power [4, 5, 8] was first conducted across all subjects to empirically derive frequency bands for analysis, resulting in LOW (4–16 Hz), BETA (17–33 Hz), and GAMMA (34–55 Hz) ranges. Next, a spatial PCA (Covariance matrix, Promax Rotation, Kappa 3 with Kaiser normalization) [4, 5, 20, 21] was completed on the broadband grand-averaged ERP waveforms (used for conventional ERP analyses) and then once for each frequency band. Figure 1 and Supplementary Figure 1 display the time courses and sPCA weights (topographies) for each waveform (ERP-voltage, LOW, BETA, and GAMMA). Weights were then multiplied by 64-sensor broadband ERP waveforms at each time point and summed across sensors, yielding four "virtual sensors" that were then plotted over time. An additional step for LOW, BETA, and GAMMA involved convolving the virtual sensor with modified Morlet wavelets (4–55 Hz, 4-ms steps, 1 cycle at lowest to 8 cycles at highest) [22-24] to derive oscillatory power waveforms for each frequency bin. For BETA, two sPCA components were derived; weighted averages of the two power waveforms were summed to derive a single waveform for analysis. This resulted in four sets of time courses that were analyzed instead of 64 separate sensors, efficiently summarizing the spatial distributions, minimizing the number of statistical

comparisons necessary, and maximizing the signal/noise ratio of the ERP data.

Prior analyses had determined no group by project interactions and highly consistent results between studies [8]. Data were binned into 10 msec segments. In order to adjust for age effects, healthy aging effects were modeled by regressing time-bin amplitudes on age for healthy subjects. When beta coefficients for age effects were significant ($p < .05$), data for all subjects within the time bin were adjusted by removing the predicted impact of age on waveform amplitude prior to group comparisons [13]. Time-bins selected for use in the bio-factor PCA analyses were based on significant group effects [11]. For each component the total mean and standard deviation across time points were calculated within each study sample (B-SNIP1 and replication) and used to standardize each time point for each subject within each study. Standardized voltage ERP, power ERP and frequency waveforms, and their 99% confidence intervals, were plotted after averaging psychosis by study and healthy participants by study (Figure 1, and Supplementary Figure 1).

Auditory Oddball Task

Analyses generally followed procedures established in Ethridge et al. [3] and Clementz et al. [11] and were applied to all subjects in this paper. Minor modifications from B-SNIP1 publications ensured standardized data quality control between projects [9]. Like with the paired-stimuli analyses, changes to the scoring procedures were primarily to frequency domain quantification and were made to simplify the analysis steps to improve the ease of replication and verification by other laboratories. The procedures described in Parker et al. [9] were used for this report.

Stimuli. Recording conditions were equivalent and stimulus presentation and recording equipment were identical across sites. Seated in a sound and electrically shielded booth (ambient sound 5 61–63 dB; luminance 5 .11–.12 foot-candle), subjects listened to tones delivered through headphones. Stimuli were 567 standard (1000 Hz) and 100 target (1500 Hz) tones presented in

pseudorandom order (1300 msec inter-trial interval). Subjects were asked to press a button when a target was detected. Subjects refrained from smoking 1 hour before testing.

Spatiotemporal Data Reduction. Data were segmented into 1000-msec epochs from 250 msec before to 750 msec after stimulus. The 250 ms pre-stimulus period was used for baseline adjustment. Epochs containing activity ±75 μV were eliminated. Data from good trials were averaged across trial types within a subject to create 64-sensor event-related potentials (ERPs).

In order to maximize use of available spatial, temporal, and oscillatory information in the evoked auditory response, a frequency-wise PCA (Covariance matrix, Promax Rotation, Kappa 3 with Kaiser normalization) of evoked power [4, 5, 9] was first conducted across all subjects to empirically derive frequency bands for analysis, resulting in LOW (1–10 Hz), BETA (11–30 Hz), and GAMMA (31–50 Hz) ranges. Next, a spatial PCA (Covariance matrix, Promax Rotation, Kappa 3 with Kaiser normalization) [4, 5, 8, 9, 20, 21] was completed on the broadband grand-averaged ERP waveforms (used for traditional ERP analyses) and then once for each frequency band for Target and Standard trials. Target trials had two components for ERP and each frequency band. Standards had one component for ERP and each frequency band (see Figure 1 and Supplementary Figure 1).

Prior analyses had determined no group by project interactions and highly consistent results between studies [9]. Data was binned into 10 ms segments. In order to minimize age effects, healthy aging effects were modeled by regressing time-bin amplitudes on age for healthy participants. When beta coefficients for age effects were significant ($p < .05$), data for all subjects within the time bin were adjusted by removing the predicted impact of age on waveform amplitude prior to group comparisons [13]. Time-bins selected for use in the PCA analyses were based on prior analyses [11]. For visualization, for each component the total mean and standard deviation across time points were calculated within each study sample (B-SNIP1 and replication) and used to standardize each time point for each subject within each study. Standardized Voltage/Power ERP and frequency waveforms

and their 99% confidence intervals were plotted after averaging psychosis by study and healthy participants by study (Figure 1 and Supplemental Figure 1).

Intrinsic EEG Activity (IEA)

Data came from the 9-10 sec inter-pair interval of the paired-stimuli task. Epochs consisted of EEG from 500 ms after the second click of each trial to 500 ms before the first click of the next trial. EEG data were pre-processed following method described above and in Thomas et al [25].

Time-frequency transformation. Data were transformed into the time-frequency (TF) domain using the following approach. In EEGLAB [19], FFTs were computed on 50% overlapping Hanning tapered windows (1-55 Hz, 1000ms steps, 1 Hz resolution) for each 9-sec inter-pair epoch, resulting in 17 time bins per epoch [500-8500 ms in 500 ms bins]. Power values (squared absolute values of complex FFT outputs) were then converted to decibels (10*log10). In order to determine the stability of the power values across time, an interclass correlation was calculated for each sensor and frequency across time bins using all participants' data. Since all ICCs were > .96, power values were then averaged over time bins. In order to capture maximum explanatory variance across variables, avoid information redundancy, and reduce the number of statistical comparisons, frequency data reduction was accomplished by principal component analysis (PCA: Covariance matrix, Promax Rotation, Kappa 3 with Kaiser normalization), see Supplementary Figure 2). The 55 frequencies were reduced to four primary bands via PCA (97% variance explained): delta/theta (1-7Hz), alpha (8-15 Hz), beta (16-30 Hz), and gamma (31-55 Hz). An additional spatial PCA (Covariance matrix, Promax Rotation, Kappa 3 with Kaiser normalization; variance range: 37-49%) was performed on each frequency band in order to reduce the data from 64 sensors to one virtual sensor [4, 5, 8, 9]. Structure matrix values for the four frequencies were: gamma=.81; delta=.88; alpha=.92; and beta=.94. Adjustments for age were as presented above.

Auditory Steady-State

Subjects listened to 166 sinusoidally amplitude modulated broadband noise stimuli at 20 (50 trials), 40 (50 trials), and 80 (50 trials) Hz and 16 unmodulated noise (duration 1500 ms; carrier pitch 1000 Hz; randomly ordered; [12]). Broadband noise bursts were used since they are known to elicit the most robust aSSRs, especially at higher frequencies [26-28]. Stimuli were presented binaurally through headphones at 75 dB SPL with an inter-trial interval of 1 sec. Subjects were instructed to count the number of unmodulated noise bursts to maintain continuous investment in the stimuli. Data were segmented into 3000-ms epochs from 750 ms pre- to 750 ms post- stimulus onset and down-sampled to 500 Hz and digitally band pass filtered from 0.5 Hz to 100 Hz (zero-phase filter; roll-off: 6 and 48 dB/octave, respectively). Only the 40 Hz trials were included in the present analyses because they optimally activate auditory cortex [27].

ERP Analysis. These procedures followed those in Parker et al. [12] exactly. Event-related potentials (ERPs) were calculated for each sensor and subject. The baseline period was defined as -100 ms prior to stimulus onset and subtracted from the grand average ERP for each subject. Eleven sensors with peak auditory response ('F1', 'Fz', 'F2', 'FC3', 'FC1', 'FCz', 'FC2', 'FC4', 'C1', 'Cz', 'C2') were identified and averaged for a final ERP. The ERPs were then standardized for each subject using the total mean and standard deviation from all subjects and time points (-100 to 350 ms). This was done in order to be comparable to ERP results from the paired-stimuli and oddball analyses. The time-period from 90-110 ms was selected as the N100 variable, and from 180-220 ms was selected for the P200 variable (see Figure 5B).

Total Power. Single-trial voltage data for each subject and sensor were converted to the time-frequency domain yielding complex numbers for points ranging from −500-2000 ms in 2 ms bins and 1 to 90 Hz following previously published methods [12, 28]. Power values (squared absolute values of

complex FFT outputs) were then converted to decibels ($10 * \log10$) and averaged across trials. The power response from sensors ('F1', 'Fz', 'F2', 'FC3', 'FC1', 'FCz', 'FC2', 'FC4', 'C1', 'Cz', 'C2') were selected and averaged together from the onset of the stimulus to the end of the steady state period (0-1500 ms) at 40 Hz (see Figure 5C).

Data Analyses

The requirement for inclusion in this project was available data on a majority of the biomarker variable classes. Data were available for the BACS from 97.5% of participants, for saccades from 86.6% of participants, for SST from 76.3% of participants, and for EEG/ERP from 83.2% of participants. Estimates of missing values were generated via a regression-based multiple imputation method [29] as implemented in SAS PROC MI using all available information from other biomarker variables. Multiple estimates from 1000 iterations were combined to provide final estimates of the missing values.

Numerical taxonomy outcomes were obtained using *k*-means clustering in SPSS. First, as described in the results, the number of clusters given the bio-factor data were determined using the gap statistic [30], which provides a formalization of the point at which within-cluster dispersion (pooled within-cluster sum of squares from the centroid) becomes less pronounced as a function of the number of clusters assumed. For our case, a null distribution was calculated by randomly shuffling bio-factors across observations (sampling with replacement) such that values for the BACS bio-factor, for instance, were randomly paired with values on all other bio-factors used for psychosis subgroups construction. We generated 1000 samples in this fashion. Plots for the null distributions (mean of the middle 99% of bootstrapped samples) and the actual data as a function of number of clusters assumed are shown in Supplementary Figure 3 for both B-SNIP1 and the replication samples. In both cases, the gap outcome indicated that the most parsimonious solution was three clusters given the bio-factor

data. We also used Two-Step pre-clustering procedure [31, 32] in SPSS. The Two-Step outcome is shown in Supplementary Table 6, which again indicates, for both B-SNIP1 and the replications samples, three clusters is the most parsimonious solution. In addition, as in Mothi et al. [33], we show the silhouette plots and values for cluster solutions from 2 to 10 for both B-SNIP1 (Supplementary Figure 6) and replication samples (Supplementary Figure 7).

For *k*-means, cases were assigned to clusters based on Euclidean distance only using available data so that imputed data would not influence the outcomes. Bio-factors were standardized to mean 0 and unit variance within B-SNIP1 and replication samples prior to clustering. There were no values outside of four standard deviations from the overall sample means that could have excessively influenced the cluster outcomes. The *k*-means algorithm was run with clusters=3, max iteration=1000, converge rate=0, and based on pairwise selection of data. See Supplementary Table 10 for the values used to standardize psychosis case values and the final clustering solutions for each Biotype. Assignment to a Biotype for every psychosis case was determined by calculating the Euclidean distance from every Biotype centroid. The smallest Euclidean distance defined Biotype membership for every case.

Canonical discriminant analyses in SPSS were used to efficiently summarize bio-factors that maximally differentiating groups (DSM diagnoses or Biotypes). Group membership was the classification variable and the bio-factors were the predictors. This analysis method eased visualization of the group differentiations, allowed a simple metric for comparing groups on multiple bio-factors simultaneously, and proved a means for calculated optimal effect size separations between subgroups (see Figures 4B and 4D). For DSM, there was only one significant function, with structure matrix values: BACS=.70; paired-stimuli S2=.55; antisaccade=-.47; P300 ERP=.46; N100=.27; SST=.19; P200 ERP=-.18; ongoing neural activity=.14; and latency=-.01. For Biotypes, there were two significant functions. Structure matrix values for the first significant function were: P300

ERP=.59; N100 ERP=.52; ongoing neural activity=.40; BACS=.38; paired-stimuli S2=.34; SST=.31; antisaccade=-.29; P200 ERP=.22; and latency=-.15. Structure matrix values for the second significant function were: antisaccade=.52; ongoing neural activity=.51; P200 ERP=.45, SST=-.41; BACS=-.29; P300 ERP=.26; latency=-.11; N100 ERP=.05; and paired-stimuli S2=-.04.

In addition, the canonical solutions using bio-factor data for both Biotypes and DSM diagnoses were used to classify cases using a jackknife (leave one out) procedure. The outcomes of the classification accuracies are shown in Supplementary Table 9. Classification accuracies were 90.9% for Biotypes, X^2(2 df)=2219.5, p<.0001, and 44.1% for DSM diagnoses, X^2(2 df)=84.6, p<.0001, both above chance. Biotypes classification accuracy was significantly better than for DSM diagnoses, X^2(2 df)=2327.9, p<.0001.

References

1. Hill SK, Reilly JL, Keefe RS, et al. Neuropsychological impairments in schizophrenia and psychotic bipolar disorder: findings from the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP) study. *Am J Psychiatry*. Nov 2013;170(11):1275-84. doi:10.1176/appi.ajp.2013.12101298
2. Reilly JL, Frankovich K, Hill S, et al. Elevated antisaccade error rate as an intermediate phenotype for psychosis across diagnostic categories. *Schizophr Bull*. Sep 2014;40(5):1011-21. doi:10.1093/schbul/sbt132
3. Ethridge LE, Soilleux M, Nakonezny PA, et al. Behavioral response inhibition in psychotic disorders: diagnostic specificity, familiality and relation to generalized cognitive deficit. *Schizophr Res*. Nov 2014;159(2-3):491-8. doi:10.1016/j.schres.2014.08.025
4. Ethridge LE, Hamm JP, Pearlson GD, et al. Event-related potential and time-frequency endophenotypes for schizophrenia and psychotic bipolar disorder. *Biol Psychiatry*. Jan 15 2015;77(2):127-36. doi:10.1016/j.biopsych.2014.03.032
5. Hamm JP, Ethridge LE, Boutros NN, et al. Diagnostic specificity and familiality of early versus late evoked potentials to auditory paired stimuli across the schizophrenia-bipolar psychosis spectrum. *Psychophysiology*. Apr 2014;51(4):348-57. doi:10.1111/psyp.12185
6. Gotra MY, Hill SK, Gershon ES, et al. Distinguishing patterns of impairment on inhibitory control and general cognitive ability among bipolar with and without psychosis, schizophrenia, and schizoaffective disorder. *Schizophr Res*. Jul 2020;doi:10.1016/j.schres.2020.06.033
7. Huang L, Jackson B, Rodrigue A, et al. Antisaccade Error Rate and Gap Effects in Psychosis Syndromes from B-SNIP2. *Psychological Medicine*. 2021;in press
8. Parker DA, Trotti RL, McDowell JE, et al. Auditory paired-stimuli responses across the psychosis and bipolar spectrum and their relationship to clinical features. *Biomarkers in Neuropsychiatry*. 2020;3:100014. doi:https://doi.org/10.1016/j.bionps.2020.100014

9. Parker D, Trotti R, McDowell J, et al. Auditory Oddball Responses Across the Schizophrenia-Bipolar Spectrum and their Relationship to Cognitive and Clinical Features. . *American Journal of Psychiatry*. 2021;in press
10. Tamminga CA, Ivleva EI, Keshavan MS, et al. Clinical phenotypes of psychosis in the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP). *Am J Psychiatry*. Nov 2013;170(11):1263-74. doi:10.1176/appi.ajp.2013.12101339
11. Clementz BA, Sweeney JA, Hamm JP, et al. Identification of Distinct Psychosis Biotypes Using Brain-Based Biomarkers. *Am J Psychiatry*. Apr 01 2016;173(4):373-84. doi:10.1176/appi.ajp.2015.14091200
12. Parker DA, Hamm JP, McDowell JE, et al. Auditory steady-state EEG response across the schizo-bipolar spectrum. *Schizophr Res*. May 9 2019;doi:10.1016/j.schres.2019.04.014
13. Dukart J, Schroeter ML, Mueller K, Initiative AsDN. Age correction in dementia--matching to a healthy brain. *PLoS One*. 2011;6(7):e22193. doi:10.1371/journal.pone.0022193
14. Keefe RS, Goldberg TE, Harvey PD, Gold JM, Poe MP, Coughenour L. The Brief Assessment of Cognition in Schizophrenia: reliability, sensitivity, and comparison with a standard neurocognitive battery. *Schizophr Res*. Jun 2004;68(2-3):283-97. doi:10.1016/j.schres.2003.09.011
15. Keefe RS, Poe M, Walker TM, Harvey PD. The relationship of the Brief Assessment of Cognition in Schizophrenia (BACS) to functional capacity and real-world functional outcome. *J Clin Exp Neuropsychol*. Feb 2006;28(2):260-9. doi:10.1080/13803390500360539
16. Keefe RS, Harvey PD, Goldberg TE, et al. Norms and standardization of the Brief Assessment of Cognition in Schizophrenia (BACS). *Schizophr Res*. Jul 2008;102(1-3):108-15. doi:10.1016/j.schres.2008.03.024
17. Hochberger WC, Hill SK, Nelson CL, et al. Unitary construct of generalized cognitive ability underlying BACS performance across psychotic disorders and in their first-degree relatives. *Schizophr Res*. Jan 2016;170(1):156-61. doi:10.1016/j.schres.2015.11.022
18. McDowell JE, Clementz BA. The effect of fixation condition manipulations on antisaccade performance in schizophrenia: studies of diagnostic specificity. *Exp Brain Res*. Jun 1997;115(2):333-44.
19. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*. Mar 2004;134(1):9-21. doi:10.1016/j.jneumeth.2003.10.009
20. Carroll CA, Kieffaber PD, Vohs JL, O'Donnell BF, Shekhar A, Hetrick WP. Contributions of spectral frequency analyses to the study of P50 ERP amplitude and suppression in bipolar disorder with or without a history of psychosis. *Bipolar Disord*. Nov 2008;10(7):776-87. doi:10.1111/j.1399-5618.2008.00622.x
21. Dien J, Khoe W, Mangun GR. Evaluation of PCA and ICA of simulated ERPs: Promax vs. Infomax rotations. *Hum Brain Mapp*. Aug 2007;28(8):742-63. doi:10.1002/hbm.20304
22. Ethridge LE, Hamm JP, Shapiro JR, et al. Neural activations during auditory oddball processing discriminating schizophrenia and psychotic bipolar disorder. *Biol Psychiatry*. Nov 1 2012;72(9):766-74. doi:10.1016/j.biopsych.2012.03.034
23. Hamm JP, Dyckman KA, McDowell JE, Clementz BA. Pre-cue fronto-occipital alpha phase and distributed cortical oscillations predict failures of cognitive control. *J Neurosci*. May 2012;32(20):7034-41. doi:10.1523/JNEUROSCI.5198-11.2012
24. Hamm JP, Ethridge LE, Shapiro JR, et al. Spatiotemporal and frequency domain analysis of auditory paired stimuli processing in schizophrenia and bipolar disorder with psychosis. *Psychophysiology*. Apr 2012;49(4):522-30. doi:10.1111/j.1469-8986.2011.01327.x

25. Thomas O, Parker D, Trotti R, et al. Intrinsic neural activity differences in psychosis biotypes: Findings from the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP) consortium. *Biomarkers in Neuropsychiatry*. 2019;1:100002. doi:https://doi.org/10.1016/j.bionps.2019.100002
26. John MS, Lins OG, Boucher BL, Picton TW. Multiple auditory steady-state responses (MASTER): stimulus and recording parameters. *Audiology*. 1998 Mar-Apr 1998;37(2):59-82. doi:10.3109/00206099809072962
27. Picton TW, John MS, Dimitrijevic A, Purcell D. Human auditory steady-state responses. *Int J Audiol*. Jun 2003;42(4):177-219. doi:10.3109/14992020309101316
28. Hamm JP, Gilmore CS, Clementz BA. Augmented gamma band auditory steady-state responses: support for NMDA hypofunction in schizophrenia. *Schizophr Res*. Jun 2012;138(1):1-7. doi:10.1016/j.schres.2012.04.003
29. Rubin DB. *Multiple imputation for nonresponse in surveys*. Wiley series in probability and mathematical statistics Applied probability and statistics,. Wiley; 1987:xxix, 258 p.
30. Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society Series B-Statistical Methodology*. 2001;63:411-423. doi:Doi 10.1111/1467-9868.00293
31. Norusis MJ. *SPSS 15.0 advanced statistical procedures companion*. Prentice Hall; 2007:xiv, 418 pages.
32. Tkaczynski A. Segmentation Using Two-Step Cluster Analysis. In: Dietrich T, Rundle-Thiele S, Kubacki K, eds. *Segmentation in Social Marketing: Process, Methods and Application*. Springer Singapore; 2017:109-125.
33. Mothi SS, Sudarshan M, Tandon N, et al. Machine learning improved classification of psychoses using clinical and biological stratification: Update from the bipolar-schizophrenia network for intermediate phenotypes (B-SNIP). *Schizophr Res*. May 25 2018;doi:10.1016/j.schres.2018.04.037